DOCUMENT RESUME

ED 106 343 TM 004 458

AUTHOR Forster, Fred
TITLE An Alternative to Ancova When Group Regressions are Heterogeneous: The Generalized Johnson-Neyman Procedure.
PUB DATE [Apr 75]
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Analysis of Covariance; Comparative Analysis; Experimental Groups; *Heterogeneous Grouping; Human Relations; *Hypothesis Testing; Models; *Multiple Regression Analysis; Research Methodology; Statistical Analysis; Teacher Education; *Tests of Significance
IDENTIFIERS *Johnson Neyman Technique

ABSTRACT

The purpose of this study was to develop the Johnson-Neyman Procedure (JN-Procedure) appropriate to multiple groups and covariables, and demonstrate its use in the analysis of group differences. A sequence of significance tests which makes it possible to identify the most parsimonious analysis of group differences appropriate to a given set of data is presented. In addition, there is an orignial mathematical derivation of the JN-Procedure appropriate to multiple groups and covariates, and a demonstration of its superiority to ANCOVA when group regressions are heterogeneous. (Author)

An Alternative to Ancova When Group Regressions

Are Heterogeneous:

The Generalized Johnson-Neyman Procedure

by

Fred Forster
Portland Public Schools
631 N.E. Clackamas Street
Portland, Oregon
97208

Presented at the 1975 Annual Meeting of the American Educational Research Association, Washington, D.C., March 30 to April 3, 1975.

In a variety of research situations it is desirable to increase the precision of the analysis of differences among treatments by using one or more covariate scores. A reoccuring (though often undetected) problem which occurs in this type of analysis is the presence of heterogeneous group regressions which indicate that the effect of the treatments is related to specific characteristics of the students. If analysis of covariance (ANCOVA) is mistakenly applied when the group regressions are heterogeneous, it can lead to erroneous conclusions since the Type II error of the analysis increases as the degree of heterogeneity increases (Peckham, 1968). An alternate procedure has been developed to analyze heterogeneous group regressions in research situations involving two groups and two covariates (Johnson and Neyman, 1936). While some efforts have been made to extend this procedure to accomodate multiple groups and covariates (Ableson, 1953; Potthoff, 1964), the practical application of this technique has been limited to two groups and one or two covariates. The problem addressed in this paper is the specification of a methodology for efficiently determining when group regressions are heterogeneous, and for applying the Johnson-Neyman procedure regardless of the number of groups or covariates in the analysis. This methodology will be discussed in relation to the analysis of an experiment in human relations training for classroom teachers.

THEORETICAL DEVELOPMENT

The approach followed here is similar to that proposed by Pearson, Neyman and their students (Pearson and Neyman, 1928; Kolodziejczyk, 1935; Johnson and Jackson, 1959). Two alternative regression models are proposed and the maximum likelihood method is used to identify the sum of squares and degrees of freedom appropriate to each model. By subtracting the smaller sum of squares from the larger and forming the appropriate F-ratio, it is possible to develop a statistical test between the models. An important aspect of this procedure is the sequencing of these statistical tests of hypotheses so that the minimum necessary effort is expended. The system of tests described below has been developed to achieve this result.

Statistical Test I - Do Any Significant Group Differences Exist

This test is intended to provide a go/no-go criteria for continuing with the analysis. If the F-ratio is not statistically significant then the analysis should be terminated, otherwise it is worthwhile to proceed to additional analyses.

The two regression models employed for this test are:

H_0: $Y_1 = \overline{B}\,\overline{X}_1$, $Y_2 = \overline{B}\,\overline{X}_2, \ldots, Y_k = \overline{B}\,\overline{X}_k$

(null hypothesis: all groups have the same regression equation) and

H_A: $Y_1 = \overline{B}_1\overline{X}_1$ $Y_2 = \overline{B}_2\overline{X}_2, \ldots, Y_k = \overline{B}_k\overline{X}_k$

(alternative hypothesis: the groups have different regression equations)

Using the maximum likelihood method (Forster, 1974) it is possible to show that the sum of squares appropriate to H_A is:

$$S_A^2 = \sum_{i=1}^{k} \sum_{j=1}^{N_i} (Y_{i,j} - \bar{X}_{i,j} \hat{B}_i)^2$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{N_i} Y_{i,j}^2 - \bar{R}_i' \hat{B}_i$$

where

$$\bar{R}_i = \sum_{j=1}^{N_i} Y_{i,j} \bar{X}_{i,j}$$

$$\bar{T}_i = \sum_{j=1}^{N_i} \bar{X}_{i,j}' \bar{X}_{i,j}$$

$$\hat{B}_i = \bar{T}_i^{-1} \bar{R}_i$$

$$df_a = \sum_{i=1}^{k} N_i - k(n+1)$$

Similarly, the sum of squares appropriate to H_o is:

$$S_1^2 = \sum_{i=1}^{k} \sum_{j=1}^{N_i} Y_{i,j}^2 - \bar{R} \bar{B}$$

where

$$R = \sum_{i=1}^{k} \bar{R}_i$$

$$T = \sum_{i=1}^{k} \bar{T}_i$$

$$\hat{B} = \bar{T}^{-1} \bar{R}$$

$$df_1 = \sum_{i=1}^{k} N_i - (n+1)$$

The appropriate F-ratio is:

$$F_1 = \frac{(S_1^2 - S_A^2) / (df_1 - df_A)}{S_A^2 / df_A}$$

Statistical Test II - Do The Groups Have Heterogeneous Regressions

The two regression models employed for this test are H_A as defined before, and

$$H_2: \; Y_1 = b_{0,1} + \overline{A}\,\overline{Z}_1; \; Y_2 = b_{0,2} + \overline{A}\,\overline{Z}_2; \; . \; . \; .; \; Y_k = b_{0,k} + \overline{A}\,\overline{Z}_k$$

(null hypothesis: group regressions are homogeneous although group means may be different)

where

$$\overline{A}_i = (b_{1,i},\, b_{2,i},\, . \; . \; . \; .\,,\, b_{n,i})$$

$$\overline{Z}_i = (X_{1,i},\, X_{2,i},\, . \; . \; . \; .\,,\, X_{n,i})$$

$b_{0,i}$ = the adjusted mean for i-th group

The sum of squares appropriate to H_2 is:

$$S_2^2 = \sum_{i=1}^{k} \sum_{j=1}^{N_i} Y_{i,j}^{\;2} - \overline{R}_H' \hat{B}_H$$

where

$$\overline{R}_H' = \left(\sum_{j=1}^{N_i} Y_{1,j},\; . \; . \; . \; ,\; \sum_{j=1}^{N_k} Y_{k,j},\; \sum_{i=1}^{K} \sum_{j=1}^{N_i} Y_{i,j}\, \overline{Z}_{i,j}\right)$$

$$\overline{T}_H = \begin{array}{cl} \left[\begin{array}{c|c} \mathrm{diag}\;(N_1, \ldots, N_k) & \begin{bmatrix} \sum_{j=1}^{N_1} Z_{1,j} \\ \cdots \\ \sum_{j=1}^{N_k} Z_{k,j} \end{bmatrix} \\ \hline \begin{bmatrix} \sum_{j=1}^{N_1} Z_{i,j} \\ \cdots \\ \sum_{j=1}^{N_k} Z_{k,j} \end{bmatrix}' & \sum_{i=1}^{K} \sum_{j=1}^{N_i} Z_{i,j}'\; Z_{i,j} \end{array}\right] & \begin{array}{l} \}\, k \\ \\ \}\, n \end{array} \\ \quad\underbrace{\qquad\qquad\qquad}_{k} \qquad \underbrace{\qquad\qquad\qquad}_{n} & \end{array}$$

$$\hat{B}_H = \bar{T}_H^{-1} \bar{R}_H$$

$$df_2 = \sum_{i=1}^{k} N_i - (k+n)$$

The appropriate F-ratio is:

$$F_2 = \frac{(S_2^2 - S_A^2) / (df_2 - df_A)}{S_A^2 / df_A}$$

The decision rule for this test is:

(1) when F_2 is statistically significant conclude group regressions are heterogeneous.

(2) when F_2 is not statistically significant use ANCOVA.

Interestingly, if ANCOVA is the indicated procedure, the necessary sums of squares are already available:

$$F_{ANCOVA} = \frac{(S_1 - S_2) / (df_1 - df_2)}{S_2 / df_2}$$

The Generalized Johnson-Neyman Procedure

When heterogeneous group regressions have been identified, it is necessary to qualify the effectiveness of a treatment on the basis of specific characteristics of subjects as reflected by the covariables.

The two regression models for this test are H_A as defined before and

$$H_{jn}: \ \overline{W}\,\overline{B}_1 = \ldots = \overline{W}\,\overline{B}_k;\ Y_1 = \overline{X}_1\overline{B}_1,\ \ldots,\ Y_k = \overline{X}_k\overline{B}_k$$

(null hypothesis: the treatments are equally effective for W, a specified set of covariate scores)

The determination of the sum of squares appropriate to H_{jn} is more involved than the previous examples. Since the second part of H_{jn} is identical to that for H_A, a similar equation results for the sum of squares:

$$S^2_{jn} = \sum_{i=1}^{K} \sum_{j=1}^{N_i} Y^2_{i,j} - \sum_{i=1}^{K} \overline{R}_i' \tilde{B}_i$$

The calculation of $\tilde{B}_i$ requires the introduction of the constraints $\overline{W}\,\overline{B}_1 = \ldots = \overline{W}\,\overline{B}_k$ through LaGrangian multipliers, resulting in the following system of equations:

$$\tilde{B}_1 = \hat{B}_1\ C_1\ \overline{W}\,\overline{T}_1^{-1}$$

$$\ldots$$

$$\tilde{B}_k = \hat{B}_k - C_k\ \overline{W}\,\overline{T}_k^{-1}$$

Returning to the basic constraint that $\overline{W}(\tilde{B}_i - \tilde{B}_j) = 0$ where $\tilde{B}_i \neq \tilde{B}_j \neq \overline{0}$ and imposing an additional constraint (for simplicity) that $\sum_{i=1}^{K} C_i = 0$ these equations can be expressed as:

$$\overline{W} \cdot \begin{bmatrix} (\hat{B}_1 - \hat{B}_2) \\ \ldots \\ (\hat{B}_{k-1} - \hat{B}_k) \\ 0 \end{bmatrix} = \overline{C} \cdot \begin{bmatrix} (W\,T_1^{-1}W - W\,T_2^{-1}W)\ldots & 0 & 0 \\ \ldots \quad \ldots & \ldots & \ldots \ \ldots \\ 0 \quad 0 & \ldots & (W\,T_{k-1}^{-1}\,W - W\,T_k^{-1}W) \\ 1 \quad 1 & \ldots & 1 \quad 1 \end{bmatrix}$$

It is possible to solve for $\bar{C} = \bar{U}^{-1} \bar{K}$ where $\bar{U}$ and $\bar{K}$ denote the right side matrix and left side vector respectively. This makes it possible to solve for $\tilde{\bar{B}}_i$ and the appropriate sum of squares:

$$S_{jn} = \sum_{i=1}^{k} \sum_{j=1}^{N_i} Y_{i,j}^2 - \sum_{i=1}^{k} \bar{R}_i' \tilde{\bar{B}}_i$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{N_i} Y_{i,j}^2 - \sum_{i=1}^{k} \bar{R}_i' (\hat{B}_i - c_i \bar{W} \bar{T}_i^{-1})$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{N_i} Y_{i,j}^2 - \sum_{i=1}^{k} \bar{R}_i' \hat{B}_i + \sum_{i=1}^{k} c_i \bar{W} \hat{B}_i$$

$$= S_a^2 + \sum_{i=1}^{k} c_i W \hat{B}_i$$

The appropriate F-ratio is:

$$F_{jn} = \frac{(S_{jn}^2 - S_A^2) / (k-1)}{S_A^2 / df_A}$$

The decision rule for this test is:

(1) When F_{jn} is statistically significant conclude that there is a significant difference among treatments for individuals with the characterisitics represented by the specified set of covariate scores.

(2) When F_{jn} is not statistically significant conclude that the treatments do not differ significantly for the individuals with the specified set of covariate scores.

The results of the generalized JN-Procedure can be used to classify each individual to the treatment group predicted to provide maximum benefit based on the specific set of covariable scores. Individuals not predicted to benefit significantly more from any of the methods can be assiged to fill out the necessary groups.

The use of these tests is summarized in Figure 1 showing the sequence of tests and the decisions to be made following each step.

APPLICATION

Project Description

The data analyzed to demonstrate the generalized JN-Procedure were drawn from a study by Bowers and Soar (Bowers and Soar, 1960). This study was designed to determine the impact of intensive laboratory training in human relations on the classroom performance of teachers. To collect data related to the goals of the study teachers completed several scales from the Minnesota Multiphasic Personality Inventory (MMPI) including the scales shown below:

(a) Hypochrondriasis scale (Hs),
(b) Depression scale (D),
(c) Hysteria scale (H),
(d) Psychopathic deviation scale (Pd),
(e) Masculine-Feminine interest scale (MF),
(f) Paranoia scale (Pa),
(g) Psychasthenia scale (Pt),
(h) Schizophrenia scale (Sc),
(i) Hypomania scale (Ma),
(j) Social introversion scale (SI),
(k) Lie scale (L),
(l) F validity scale (F),
(m) K validity scale (K),
(n) Anxiety scale (A),
(o) Repression scale (R),
(p) Ego-strength scale (Es),
(q) Hostility scale (Ho),
(r) Pharisaic-virtue scale (Pv).

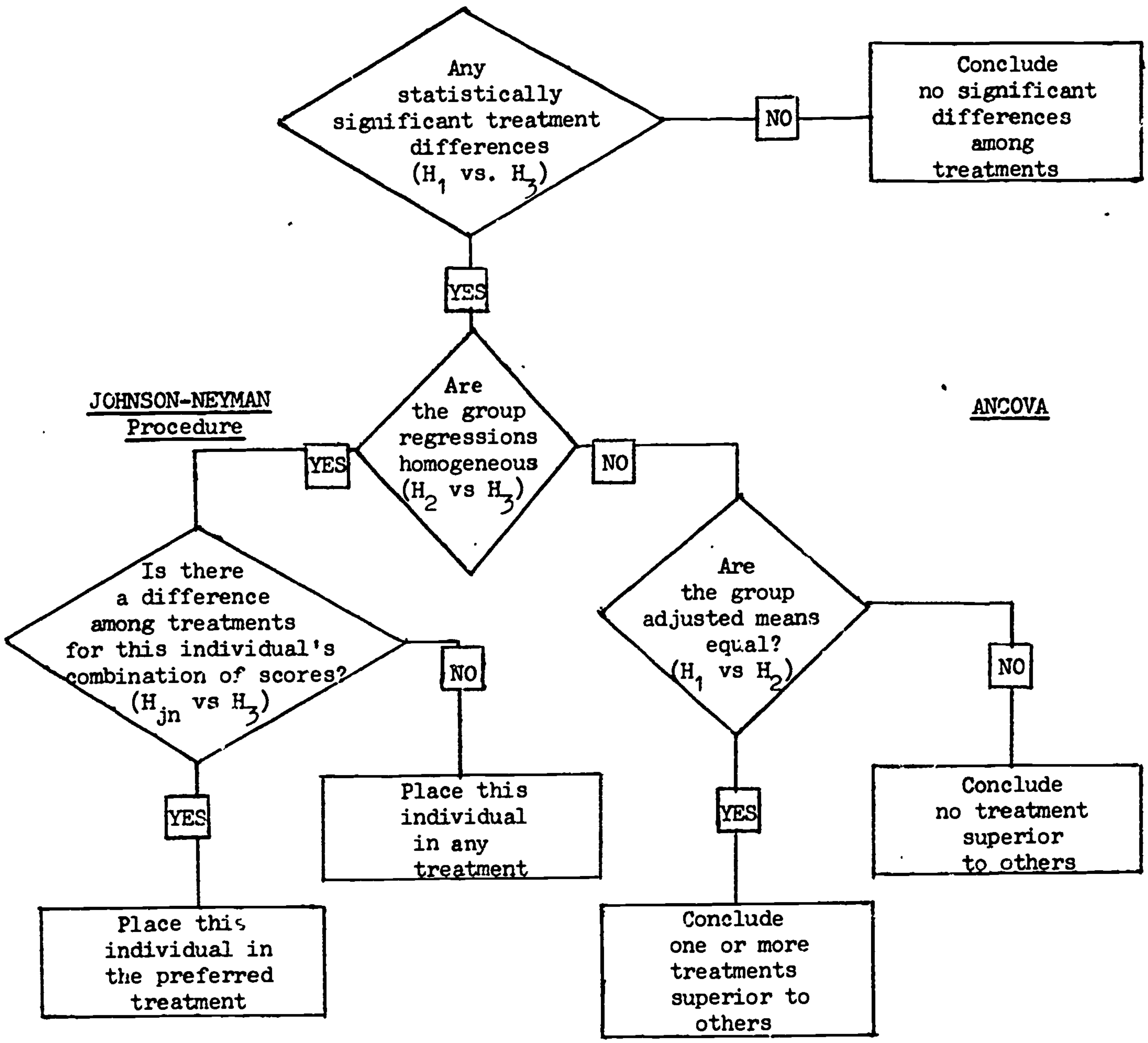

FIGURE 1
Sequence of Tests for the
Generalized Johnson-Neyman Procedure

In addition, teachers were asked to report on the number of activities which they planned cooperatively with other teachers (SR), the Observation Schedule and Record (OSCAR2E) was used to assess the social structures observed in each teachers classroom on a pre and post treatment basis (SS_{pre}, SS_{post}), and the Russell Sage Social Relations observation scale was used in each classroom. Sixty-two elementary school teachers were recruited to participate in the Bowers and Soar project. Of these, eight dropped out of the study and three others taught special education classes and were not included in the analysis. This left a total of fifty-one teachers in the study, twenty-eight in the control group, and twenty-three in the experimental group. Since participation in the experimental group required attendance at a three-week workshop either immediately after school adjourned for the summer or immediately before school opened in September, some teachers were not able to participate in either workshop. Consequently, assignment to the treatments was non-random. The experimental group was composed entirely of teachers able and willing to participate in the workshop, while the control group included some teachers unable or unwilling to participate in the workshop. In spite of this, the comparison of means and distributions for personal history data, pretraining measures, and school socioeconomic data as reported by Bowers and Soar did not reveal any significant differences between the groups. (Bowers and Soar, 1960, pp. 49-50)

During the spring semester 1959, each teacher was asked to complete inventory and questionnaire data. In addition, the OSCAR2E was used during the spring semester to complete the pretraining data.

During the summer, teachers in the experimental group attended one of two workshops on Human Relations Training. These workshops included three primary activities:

(1) Theory sessions, using a relatively traditional lecture format,

(2) skill practice, using role play, group discussion, and related activities, and

(3) training group, using unstructured group discussion.

Reading lists and a collection of reading materials were made available to participants, but no assignments were made. The following spring semester follow-up data was gathered of approximately the same nature as the pretraining data. Although a number of hypotheses were presented by Bowers and Soar, this reanalysis of their data focused on the following hypothesis proposed in their original study:

> There are no differences between teachers who have experienced intensive laboratory training in human relations and teachers who have not experienced such training with respect to the Social Structures of the classroom.

Analyses of Data

Since regression analyses have been shown to be adversely affected by heterogeneity of variance among groups with respect to the covariables used in the analysis, high intercorrelations among the covariables, and departure from a linear relationship between the dependent variable and the covariable, preliminary tests were performed to determine whether any or all of these conditions might exist. These tests indicated that there were no major departures from homogeneity of variance between the groups when the small sample sizes were taken into account. In addition, the similarity of values for the means and medians indicated that the distributions of the variables were not skewed significantly.

The data were screened to identify pairs of covariables which were so highly correlated with the SS posttraining scores as to introduce instability into the group regression equations. Using a test proposed by Lord and Novick (Lord and Novick, pp 266-269), it was decided to drop several covariables from the analysis includ k c, Sc, Pv, SI, Hs, R, D, Es and Ma scales of the MMPI.

Scatterplots were developed for each group contrasting each potential covariable with SS_{post} to identify departure from linearity in the data. The visual analysis of these graphs did not suggest a significant departure from linearity for any of the covariables. A supplementary analysis was performed to statistically test the degree to which the relationship between each covariable and the criterion significantly departed from linearity by stepwise adding terms to the regression model:

$$y = b_o + b_1 x + b_2 x^2 + \dots + b_n x^n$$

At each step the calculation was made of F-reg indicating the statistical significance of the fit and of F-dep indicating the departure from fit. The results from this procedure also indicated there were no significant departures from linearity for the covariables with SS_{post}.

The final screening procedure used was a forward stepwise regression analysis for each group, to identify the set of covariables which maximized the size of the multiple correlation coefficient where SS_{post} was the criterion variable. The sets of maximally effective covariables identified for the experimental and control groups had only two members in common, SS_{pre} and Ho. Four additional covariables were effective for either one treatment or the other and were retained in the analysis. The six covariables selected for use in this analysis were SS_{pre}, RSSR, SR, L, Pa, and Ho.

The first test contrasted the null hypothesis that no treatment differences exist with respect to means or regressions with the hypothesis that the groups differ with respect to means and regressions (H_1 vs. H_3). As shown in Table 1-A, the F-ratio attained for this test was statistically significant ($p < .01$), indicating that the analysis should proceed.

The F-ratio was statistically significant ($p < .05$) for the test of the hypothesis that there are no group differences with respect to regressions versus the hypothesis that the groups differ with respect to means and regressions (H_2 vs. H_3). This result indicated that the group regressions were heterogeneous and the JN-Procedure was the appropriate analysis for these data.

Table 2 summarizes F_{jn} for each teacher in the study for whom F_{jn} was statistically significant. As discussed in the previous analysis, the teachers with scores yielding statistically significant F-ratios fall within a region of significance, and the remaining teachers have scores which fall within a region of non-significance. In comparing the covariable scores for the teachers for whom F_{jn} is statistically significant and the regression coefficients for the experimental and control groups (as shown in Tables 1-B and 2) the effectiveness of the human relations training appears to be associated with low pretraining scores on Pa and L. This relationship is dramatized by the scores for the two teachers in the experimental group (E15 and E22) for whom the treatment is predicted to be inferior to the control condition. Those teachers had the highest scores on the RSSR and SR, two of the three highest scores scores on Ho, the lowest scores on Pa, and

TABLE 1-A

F Tests For The Generalized Johnson-Neyman Procedure Using The Bowers-Soar Data

Number of Groups = 2
Number of Covariables = 6

Test	Source	Degrees of Freedom	Sums of Squares	Mean Square	F-ratio
H_1 vs. H_3	$S_1^2 - S_3^2$	7	391.613	55.944	3.471**
H_2 vs. H_3	$S_2^2 - S_3^2$	6	278.634	46.439	2.881*
H_3	S_3^2	37	596.430	13.417	---

TABLE 1-B

Unrestricted Group Regression Coefficients

Group	Mean	Covariable					
		SS_{pre}	SR	L	RSSR	Pa	Ho
Experimental	74.416	.262	-.359	-.023	-.156	-.[illegible]59	-.139
Control	55.464	.241	-.074	-.227	.096	-.314	.233

TABLE 2

SUMMARY OF THE STATISTICALLY SIGNIFICANT F-RATIOS FOR THE JN-HYPOTHESIS

Group Membership[1]	F_{jn}	SS_{post} Predicted		Observed Scores						
		Experimental	Control	SS_{post}	SS_{pre}	RSSR	SR	Pa	L	Ho
C4	15.04**	57.720	45.985	50	42	44	39	53	56	36
C6	5.19*	57.294	51.220	54	46	48	39	53	43	40
C12	12.73**	59.474	47.492	48	42	41	33	56	50	40
C16	14.24**	54.303	47.578	50	45	48	48	53	53	38
C18	20.56**	59.624	44.406	43	45	35	39	59	53	35
C22	8.53**	52.604	40.120	38	32	37	51	56	53	29
C23	13.51**	52.492	42.010	42	44	53	54	59	56	26
C24	12.56**	59.285	48.610	49	53	41	44	56	46	33
C25	8.57**	52.679	44.181	44	46	58	51	56	63	33
E4	8.51**	54.072	45.035	48	53	51	55	62	50	29
E5	3.99*	54.081	50.030	59	47	56	44	56	50	43
E10	6.32**	52.169	44.573	54	48	58	51	65	53	35
E15	3.99*	47.741	55.528	50	57	60	68	44	50	46
E17	5.57*	53.796	48.323	54	48	46	48	59	52	47
E20	15.34**	55.536	47.372	52	43	49	42	56	50	35
E22	7.68**	48.113	61.419	46	50	68	54	41	40	57

[1] C denotes a teacher in the control group; E denotes a teacher in the experimental group; the number denotes relative position in the Bowers and Soar data tables.

*$p<.05$
**$p<.01$

relatively low scores on L. Based on these data, the experimental treatment appears to have the greatest effectiveness with teachers who have pretraining scores on the MMPI scales related to the need for acceptance and approval and paranoia, and who report a low level of student interaction in the classroom, for whom few instances of student interaction are observed during the RSSR tasks, and who are low in hostility. This conclusion is consistent with expectations that the effectiveness of the treatment would be greater for teachers who need help with regard to the level of student interaction in the classroom, who are not hostile, and who demonstrate a need for acceptance and approval.

The application of ANCOVA to these data yields an F-ratio of 5.552 with one and forty-three degrees of freedom (p .05), an adjusted mean for the experimental group of 56.337 and an adjusted mean for the control group of 53.884. Thus, the null hypothesis would be rejected by both ANCOVA and the JN-Procedure. As with the research situation, the JN-Procedure provides more information than ANCOVA concerning the specific individuals most likely to benefit from the treatment. The JN-Procedure indicates that the treatment would significantly benefit only fourteen of the fifty-one teachers in the study (27.5 per cent).

SUMMARY

The application of the Generalized Johnson-Neyman procedure to the data previously reported by Bowers and Soar suggests that the training in human relations had its most significant impact for those teachers who reported few cooperative activities with their colleagues, who demonstrated a need for acceptance, who were low in hostility and whose students did not perform well on cooperative projects. The principle advantage of the JN-Procedure over ANCOVA in this situation was that it identified the relatively few teachers who would be expected to significantly benefit from the treatment. This characteristic of the Generalized JN-Procedure has significant implications for those researchers and administrators who desire to allocate teacher training resources in the most effective way.

FF:dk

REFERENCES

Abelson, R.P. A Note on the Johnson-Neyman Technique. Psychometrika, XVII 1953, pp. 213-218.

Bowers, N.E. and Soar, R.S. Evaluation of Laboratory Human Relations Training for Classroom Teachers. USOE Cooperative Research Project No. 469. Chapel Hill: University of North Carolina, 1961.

Forster, F. The Extension of the Johnson-Neyman Procedure for the Analysis of Group Differences, Unpublished PhD Dissertation. Evanston: Northwestern University, 1974.

Johnson, P.O. and Jackson, R.W.B. Modern Statistical Methods. Chicago: Rand McNally, 1959.

Johnson, P.O. and Neyman, J. Tests of Certain Linear Hypothesis and Their Application to Some Educational Problems. Statistical Research Memoirs, I 1936, pp. 57-93.

Kolodziejczyk, S. On an Important Class of Statistical Hypotheses. Biometrika, XXVII 1935, pp. 161-190.

Neyman, J. and Pearson, E.S. On the Use and Interpretation of Certain Test Criteria for the Purpose of Statistical Inference. Biometrika, XX 1928, pp. 175-240.

Peckham, P.D. An Investigation of the Effects of Non-Homogeneous Regression Slopes Upon the F-Test of Analysis of Covariance. Laboratory of Educational Research Report, No. 16. Boulder, Colorado: University of Colorado, 1968.

Potthoff, R.F. On the Johnson-Neyman Technique and Some Extensions Thereof. Psychometrika, XXIX 1964, pp. 214-256.

FF:dk

REFERENCES

Abelson, R.P. A Note on the Johnson-Neyman Technique. Psychometrika, XVII 1953, pp. 213-218.

Bowers, N.E. and Soar, R.S. Evaluation of Laboratory Human Relations Training for Classroom Teachers. USOE Cooperative Research Project No. 469. Chapel Hill: University of North Carolina, 1961.

Forster, F. The Extension of the Johnson-Neyman Procedure for the Analysis of Group Differences, Unpublished PhD Dissertation. Evanston: Northwestern University, 1974.

Johnson, P.O. and Jackson, R.W.B. Modern Statistical Methods. Chicago: Rand McNally, 1959.

Johnson, P.O. and Neyman, J. Tests of Certain Linear Hypothesis and Their Application to Some Educational Problems. Statistical Research Memoirs, I 1936, pp. 57-93.

Kolodziejczyk, S. On an Important Class of Statistical Hypotheses. Biometrika, XXVII 1935, pp. 161-190.

Neyman, J. and Pearson, E.S. On the Use and Interpretation of Certain Test Criteria for the Purpose of Statistical Inference. Biometrika, XX 1928, pp. 175-240.

Peckham, P.D. An Investigation of the Effects of Non-Homogeneous Regression Slopes Upon the F-Test of Analysis of Covariance. Laboratory of Educational Research Report, No. 16. Boulder, Colorado: University of Colorado, 1968.

Potthoff, R.F. On the Johnson-Neyman Technique and Some Extensions Thereof. Psychometrika, XXIX 1964, pp. 214-256.

FF:dk